# This Page Is Inserted by IFW Operations and is not a part of the Official Record

### BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

#### IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents will not correct images, please do not report the images to the Problem Image Mailbox.



## 证明

#### 本证明之附件是向本局提交的下列专利申请副本

申 请 日: 2002 08 07

申 请 号: 02 1 34545.7

申请类别: 发明

发明创造名称: 文字图表等一般载体上的信息电子化的方法及其系统

申 请 人: 深圳矽感科技有限公司; 深圳市和诚科技实业有限公司

发明人或设计人:张伟



2003 年 7 月 31 日





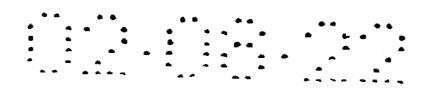
#### 权利要求书

- 1. 一种文字图表等一般载体上的信息电子化的方法, 其特征在于包括如下步骤:第一步,将应用程序中生成的 应用电子化文档编码产生二维条码电子化文档;第二步, 将所述二维条码电子化文档和所述应用电子化文档附着在 一般载体上形成纸质电子化文档;第三步,通过扫描所述 纸质电子化文档中的二维条码码图获得数字信息;第四步, 对所述数字信息进行二维条码解码,还原出所述应用电子 化文档。
- 2. 根据权利要求 1 所述的文字图表等一般载体上的信息电子化的方法,其特征在于所述第一步中产生的二维条码电子化文档中增加齿型定位系统: 所述第二步中形成的纸质电子化文档中的二维条码码图两边具有由等距离的黑白间隔构成的定位齿; 所述第三步中扫描所述二维条的四图获得的数字信息根据所述定位齿还原垂直于扫描行的内上的拉伸和压缩变形, 舍弃多出的扫描行数据, 补足缺少的扫描行数据。
- 3. 根据权利要求 1 所述的文字图表等一般载体上的信息电子化的方法,其特征在于所述第一步中产生的出条码电子化文档中增加扫描定位线: 所述第二步中的成质电子化文档的二维条码码图的一侧、或者码图的短行斜线构成的扫描定位线加于二维条码码图的数字信息。

- 4. 根据权利要求 1、2 或 3 所述的文字图表等一般载体上的信息电子化的方法, 其特征在于所述第二步中获得的纸质电子化文档中的二维条码码图具有对应被编码的应用电子化文档的不同内容或不同章节的目视可分的区块。
- 5. 根据权利要求 2 所述的文字图表等一般载体上的信息电子化的方法,其特征在于根据扫描分辨率的要求在垂直于扫描行的方向上的扫描数据拉伸时以平均间隔舍弃多出的扫描行数据,在垂直于扫描行的方向上的扫描数据压缩时以平均间隔复制临近的上一行扫描行数据插入补足。
- 6. 根据权利要求 3 所述的文字图表等一般载体上的信息电子化的方法, 其特征在于每一个扫描行与一根扫描定位线交叉。
- 8. 根据权利要求 7 所述的文字图表等一般载体上的信息电子化的系统, 其特征在于, 所述二维条码编码模块

和二维条码解码模块安装在所述电脑主机中;或者所述二维条码编码模块安装在所述印刷机或曝光机中,所述二维条码解码模块安装在所述扫描仪中。

- 9. 根据权利要求 7 或 8 所述的文字图表等一般载体上的信息电子化的系统,其特征在于: 所述扫描仪包括信号处理芯片(10)和图像传感器(20): 所述信号处理芯片(10)根据位于二维条码码图的一侧、或者位于码图的两侧、或者以不同于二维条码码图的颜色选加于二维条码码图之上的、由等间距平行斜线构成的扫描定位线获得光电探测器与被扫描文件相对移动的距离,并根据该距离控制图像传感器(20)以设定的扫描分辨率获得二维条码码图的数字信息。
- 10. 根据权利要求 7 或 8 所述的文字图表等一般载体上的信息电子化的系统, 其特征在于: 所述印刷机或曝光机是打印机, 所述打印机包括打印控制程序, 根据用户设定的打印纸质的不同分别选择最优的文本打印分辨率用于文本打印、选择最优的图形打印分辨率用于图形打印、和选择最优的二维条码打印分辨率用于二维条码码图的打印。
- 11. 根据权利要求 9 所述的文字图表等一般载体上的信息电子化的系统, 其特征在于, 所述信号处理芯片(10)根据位于二维条码码图两边的、由等距离的黑白间隔构成的定位齿, 还原垂直于扫描行的方向上的拉伸和压缩变形, 舍弃多出的扫描行数据, 插入补足缺少的扫描行数据。
- 12. 根据权利要求 9 所述的文字图表等一般载体上的信息电子化的系统, 其特征在于, 所述扫描仪还包括闪存





模块(40)、模数转换模块(50)、RAM存储模块(60)和数据接口(80): 所述信号处理芯片(10)产生控制各模块的控制信号,并对扫描输入的二进制数据进行解码,将解码后的数据存储在RAM存储模块(60)中,并根据需要输出到数据接口(80): 图像传感器(20)根据信号处理芯片(10)送来的时钟信号(Clock)和曝光同步信号(SP), 捕捉来自被扫描文件的反射光产生模拟电信号(Vout), 再输入模数转换模块(50)转变为数字信号: 所述闪存模块(40)存放信号处理芯片(10)运行所需要的程序。

13. 根据权利要求 10 所述的文字图 表等一般载体上的信息电子化的系统, 其特征在于, 打印控制程序可以根据需要访问打印数据库, 所述打印数据库中包括: 与纸张类型相应的文本打印分辨率、图形打印分辨率和二维条形码打印分辨率。



#### 说明书

### 文字图表等一般载体上的信息电子化 的方法及其系统

技术领域 本发明涉及文字图表等一般载体上的信息电子化的方法及其系统,具体地说涉及将一般载体上附着的便于人们阅读的信息快速无误地电子化的方法以及将该信息电子化输入、输出的装置的改进。

背景技术 当前人们在将印刷在纸制媒介上的文档电 子化的时候使用的方法都是将文字、图片和表格等文件扫 描到计算机上以图片格式保存,这种方法获得的图形文件 占用控件较大, 而且不方便检索、浏览以及二次排版。 了解决这个问题,有些系统就外挂了 OCR (光学字符识别 系统)来将扫描获得的图片文件转换成纯文本格式的文件, 这样处理后的文件不仅占用控件小,而且能够很方便的对 其进行检索、浏览以及二次排版。但是,这种方法存在许 多问题,例如识别率问题:OCR产品经过多年开发目前仍 然没有一家公司的产品能保证达到使人们满意的识别率, 自动识别后的文档无法保证与原文档相同。因为识别率无 法保障 100%, 所以在做完 OCR 处理后还需要人为的对识 别后的文件进行检查及修改,而且改动量比较大,尤其是 一些扫描不太清晰的文档,即使使用最好的OCR产品,识 别率都会低于30%,这时检查及修改的工作量甚至大于重 新输入的工作量。另外OCR产品都是针对指定的语言来进 行识别的,一个很好的 OCR产品也只能支持有限的几门语 言。更为重要的是扫描件在经过OCR处理之后生成纯文本 文件丢失了原文档的某些信息,例如一份打印出的 Word 文 档、除了文字内容所传达的信息以外、还包括字体信息、 表格信息、批注信息等,扫描后经 OCR 处理生成的文件不

再包含这些信息。使用 OCR 对扫描件进行识别,输入设备只能是扫描仪,没有其他可选输入设备。每输入一张纸上的文字内容就需要扫描整份文档,因此用户要花很多时间和精力。

在信息化社会里,人们需要快速将文字图表等一般载体上的信息电子化的方法及其简便宜用的系统。

技术方案 本发明的目的在于提出一种文字图表等一般载体上的信息电子化的方法及其系统,便于人们随时电子化摘录、存储、引用所阅读的书籍、报纸、杂志上的有用信息,使纸质文档和电子文档结合成为一体,便于阅读、携带和存档。

本发明公开了一种文字图表等一般载体上的信息电子化的方法,包括如下步骤:第一步,将应用程序中生成的中生对组编码产生二维条码电子化文档;第二步都附述二维条码电子化文档和所述应用电子化文档和所述应用电子化文档描述,通过扫描所述的电子化文档中的二维条码码对表得数字信息:第四步,对所述数字信息进行二维条码解码,还原出所述应用电子化文档。

也可以以不同于二维条码码图的颜色选加于二维条码码图之上,所述第三步中扫描所述二维条码码图时根据扫描定位线获得光电探测器与被扫描文件相对移动的距离,并根据该距离控制扫描仪以设定的扫描分辨率获得二维条码码图的数字信息。

本发明在应用文档生成时就将其文档信息转换成二维

附图说明 本发明包括如下附图:

图 1 是制作纸质电子化文档的流程示意图:

图 2 是纸质电子化文档电子化的流程示意图;

图 3 是通过软件方法实现文字图表等一般载体上的信息电子化的示意图:

图 4 是通过打印机编码制作纸质电子化文档的流程示意图;





图 5 是纸质电子化文档通过扫描仪解码完成电子化的流程示意图:

图 6 是可高效打印纸质电子化文档的打印机驱动程序;

图 7 是有扫描定位标志的二维条码;

图 8 是有扫描定位线的二维条码;

图 9 是有另一种扫描定位线的二维条码:

图 10 是利用扫描定位线控制扫描的原理示意图;

图 11 是具有目视可分的区域的二维条码示意图;

图 12 是一种二维条码扫描仪的电路原理框图。

具体实施例 下面结合附图并通过具体实施例对本发明做进一步详细说明。

如图 1 所示是制作纸质电子化文档的流程示意图,在各种电脑应用程序中生成的电子化文档,包括纯文本文件、图形文件、工程图纸和 Word 文档,经过二维条码生成器产生二维条码,该二维条码和在应用程序中生成的原电子化文档再通过普通打印机在打印纸上打印出原文档和相应的二维条码,生成纸质电子化文档。二维条码码图在文档中的具体位置不限于图中所示,可以由用户选择。二维条码码图的外形也不限于矩形块,还可以构成其他平面图形、例如三角形、圆形或其组合。

如图 2 所示是纸质电子化文档电子化的流程示意图, 采用扫描仪或二维条码识别器扫描纸质电子化文档中的二维条码, 再将扫描获得的数字信息送入二维条码解码器进行解码, 还原出各种原文档, 包括纯文本文件、图形文件、 工程图纸和 Word 文档等。

如图 3 所示是通过软件方法实现文字图表等一般载体上的信息电子化的示意图,采用安装在电脑主机中的应用软件层次上的二维条码生成程序和二维条码解码程序,和打印机、扫描仪进行无缝连接实现纸质电子化文档的过程。

图 3 中二维条码生成程序提供了一个文件浏览器,此

浏览器能自动支持绝大部分常用格式文档(如:Office 文件,WPS文件,CAD文件,JPG,GIF,TIF,BMP等)。对于所支持的格式文档,用户可以在此浏览器中直接打开印度,并且可以通过打印度。对于此浏览器不支持的格式文档,用户也可以在现所得。对于此浏览器不支持的格式文档,用户也可以在的打开,虽然看不见最终显示的效果,同样用户也可以在自免底板中自由选择并拖动条形码的位置然后打印,或者让条形码打印在第二页。

图 3 中二维条码解码程序提供一个操作界面,用户可以在此界面直接选择已安装的扫描仪对文档进行扫描,扫描后系统会自动找出二维码部分并进行解码,解码后的文件格式如果是内置浏览器所支持的文件格式就会直接在浏览器中显示,用户可以选择将解码后的文档保存到各种存储介质上,也可以直接打印。

本发明可以采用各种二维条码,例如 PDF 417、DATASTRIP,等等。所采用的二维条码可以根据需要插入各种纠错码。这些二维条码和纠错码都是本领域技术人员熟知的现有技术。本发明也可以采用新发明的各种二维条码和纠错码。

本发明所称的文字图表等一般载体可以是纸张等不透明的印刷载体,也可以是胶片等透明的图文载体。本发明所称的纸质电子化文档是指同时附着有易于直接阅读、辨认的文字、图形信息和对应上述信息的二维条码码图的纸张或胶片。

如图 4 所示是通过打印机编码制作纸质电子化文档的流程示意图,通过与打印机硬件的绑定,将编码器加入到打印机驱动程序中,用户可以在自己的应用程序里完成打印。打印时,会弹出一个打印界面,在控制纸张大小、打印份数的同时,选择是否打印二维条形码。



可以采用专用的条码打印机,这种打印机以热转换方式打印效率比较高,但该设备及耗材成本均较高;也可以采用普通打印机例如激光打印机打印。本发明的这种文字图表等一般载体上的信息电子化的方法原则上可以使用普通打印机实现,但是如果对普通打印机进行改进,可以获得更高的打印效率和更好的效果。

如图 5 所示是纸质电子化文档通过扫描仪解码完成电子化的流程示意图:通过与扫描仪硬件的绑定,将解码器加入到扫描仪驱动程序当中,使用户可以在自己的应用程序里扫描并自动解码。用户在扫描时系统会自动去找二维条形码,如果找到了就会提示用户选择是否还原出原文件。

使用二维条码识别器读入二维条码能大大提高输入速度,它的实现方式跟如图 3 所示的应用软件层次上的解决方案相似。

本发明对打印机的改进包括:

根据不同的数据类型选择最优打印分辨率,根据纸质的不同区别处理。打印机根据用户选择的纸质类型、编码方式和打印区域,自动调整文本数据、图形数据、二维码区域的打印分辨率,例如在好的纸张上可以打印更高分辨率的二维码而不会导致识别的误码率增加,在非是高。

如图 6 所示是本发明的一个实施例,一种打印机的控制流程。该打印机可以根据需要访问数据库,数据库中所存储条目包括:与纸张类型相应的打印分辨率、文本打印分辨率、图形打印分辨率和二维条形码打印分辨率。设 R1 为适合文本数据的打印分辨率; R2 为适合图形数据的打印分辨率; R3 为适合二维条形码的打印分辨率。为了在提高打印速度的前提下不降低打印质量,不同的类型的纸张所对应的 R1、R2 和 R3 不同。

在打印开始后,打印机控制程序根据用户选择的纸质类型访问数据库,并根据数据库内容设定相应的 R1、R2和 R3的数据,再根据需要打印的数据类型的设置当前分辨率,然后根据需要打印的数据长度完成该数据类型的打印,根据打印数据的需要转换打印分辨率继续打印,直至所有数据打印完成。

本发明的这种文字图表等一般载体上的信息电子化的方法可以采用手持扫描设备,由于二维条码是靠图像与原图面积来编码和解码,二维条码的扫描获得图像与原图一致才能保证解码正确,如果通过硬件解决这种通图,使扫描设备成本大大的增加,特别体现在手持扫描设备上。符本发明在二维条码两侧再加一组定位符号,根据变形,在号对扫描获得的图像进行处理,校正图像的拉伸变形,在解码的先行还原成精确图像,保证扫描图像与原图一致。

如图 7 所示是有扫描定位标志的二维条码。在二维条码图像两边增加齿型定位系统,利用这些"定位齿"可离图像垂直于扫描方向的拉伸形变,定位齿是等距离的照图像垂直于扫描行动,并结束于定位齿,正确的扫描发展在每一行的开始、和结束处具有规律的 01 分条。他没扫描仪在某一分辨率下一个定位齿的距离上有多进场的扫描行。当出现在中的电离上有多进场的扫描行,此时可选择均匀舍弃多出的时一个定位齿的距离上有多大个原图像,消除图像的拉伸。当出现压缩时一个定位齿陷的图像。

扫描定位标志还可以直接控制扫描仪正确采集信号,如图 8 和图 9 所示是本发明另一种有扫描定位标志的二维条码示意图,采用扫描定位线。扫描定位线为等间距平行斜线,可以位于二维条码码图的一侧,如图 8 扫描定位线位于二维条码码图左侧,当然扫描定位线也可以位于二维条



码码图右侧,或者以不同于二维条码的颜色选加于二维条码码图之上(如图 9),也可以位于二维条码码图的两侧。当扫描定位线以不同于二维条码的颜色选加于二维条码码图之上时,扫描仪可选择于扫描定位线相同颜色的光源用于获取扫描定位线的位置信息,选择于扫描定位线相反颜色的光源用于获取二维条码码图的信息。

最通用的光学扫描器包括平板扫描仪和单张馈纸扫描 仪, 在扫描时, 扫描仪的图像传感器接受来自于被扫描文 档的反射光, 图像传感器可以是多个光电二极管、光电晶 体管或电荷耦合器件(CCD),总称为光电探测器,它们对入 射光敏感并产生电子信号,这些信号被称为各个光电探测 器的像素或电荷信号。光电探测器与被扫描文档相对运动, 一次获取一行的扫描信号。在不同的扫描行,光电探测器 扫描到扫描定位线的位置不同。如图 10 所示是利用扫描定 位线控制扫描的原理示意图。假设在扫描行A在像素B接 受到扫描定位线信号,在扫描行(A+n)在像素(B+m)接受到 扫描定位线信号,则像素 B 和像素(B+m)之间的距离乘以扫 描定位线的斜率即可获得光电探测器与被扫描文档的相对 移动的距离。在扫描仪的控制电路中用积分器读取图 10A 中两个峰值的差值,当这个差值达到扫描分辨率的要求时, 启动光电探测器接受来自于被扫描文档的反射光并获得一 行扫描信号。例如,当设置扫描分辨率为 300dpi时,通过 扫描定位线获得三百分之一英寸的距离时即控制扫描仪获 取一行扫描数据。

传统二维条码的码图为一个整体,不能人为选择其中的一段进行扫描解码,本发明的二维条码的码图具有目视可分的区块,这种目视可分的区块对应被编码的原文档中的不同内容或不同章节,选择某些区块扫描解码,就相当于摘录被编码的原文档中的相应内容或相应章节。如图 12 所示是具有目视可分的区域的二维条码示意图。

如图 12 所示是一种二维条码扫描仪的电路原理框图,包括信号处理芯片 10,图像传感器 20,卡片检测模块 30,闪存模块 40,模数转换模块 50,RAM 60,调试端口 70,数据接口 80。

信号处理芯片 10 是扫描仪的核心部分,它负责产生各个模块的控制信号;对扫描输入的二进制数据进行解码;把处理后的数据存储在 RAM 60 中并根据需要输出到数据接口 80 。

图像传感器 20 根据信号处理芯片 10 送来的时钟信号 (Clock)和曝光同步信号(SP), 捕捉来自被扫描文件的反射光产生模拟电信号(Vout)。

卡片检测模块 30 检测是否有被扫描文件插入,并产生一个"Paper"信号,信号处理芯片 10 根据"Paper"信号决定是否启动或停止图像传感器 20 扫描。

闪存模块 40 存放信号处理芯片 10 运行所需要的程序。模数转换模块 50 将扫描获得的模拟信号转变为数字信号。

RAM 60 存放解码后的条码信息。

调试端口70用于程序的下载及仿真板的连接。

数据接口 80 用于把 RAM 60 内的数据输出到其它装置如电脑主机上。

本发明所称的电脑主机包括各种品牌和型号的家用、商用电脑、掌上电脑, 也包括各种有文字处理能力、或者图形处理能力的数字处理系统。

本发明公开的一种文字图表等一般载体上的信息电子化的方法和系统在很多行业具有非常广泛的应用前景,包括政府机关,行政单位,邮政、运输行业,零售业,制造业,医疗业,建筑业,个人用户等。这种文字图表等一般载体上的信息电子化的方法的广泛使用,能有效改善人们的生活,实现电脑之外的所见即所得。例如某人在早餐桌



上对所读报纸中的一个新闻感兴趣、或者坐飞机时在所读杂志上看到一篇好文章即可利用扫描笔随手摘录下来存储在 PDA 中。

政府部门下发的各种文件如果是纸质电子化文档,就可以方便准确的被引用,既节省了反复录入所需要的人力,不从网络上发放还可以避免网络安全问题导致的失密。

到银行开户时需要填写一系列表格,银行工作人员将新用户表格上的信息再输入到电脑中去。采用本发明银守图表等一般载体上的信息电子化的方法新用户先在带身的网站上面将电请表格填好,打印出申请单,然后带头的份证拿到银行去开户,银行工作人员可以利用二维条份份证拿到银行去开户,银行工作人员可以利用二维条分别器快速的将申请单上的信息还原到电脑里面。节省了大量的时间。

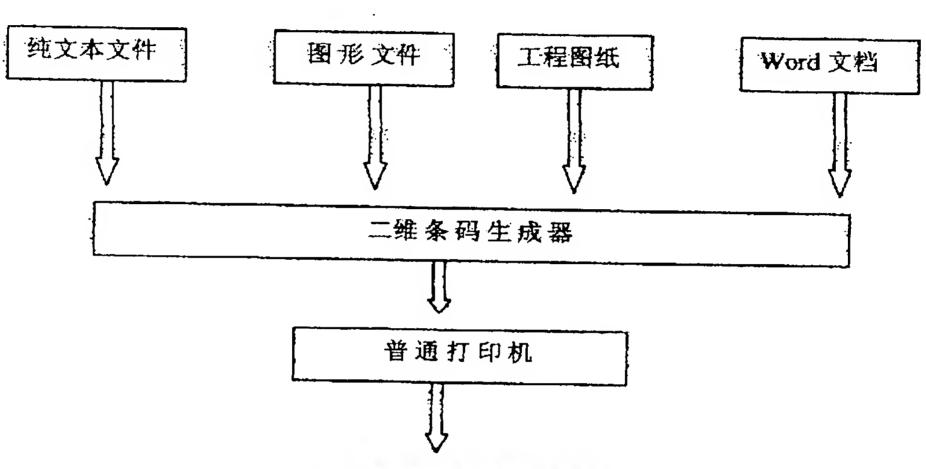
邮局、快递公司等在对信件、包裹进行分发时需要将信封、包裹标签上的内容输入电脑,然后再进行分发。手

工输入的工作量是巨大的,输入的错误发生率也相当高。而且经常由于一、两个字写的比较潦草而将发后件发布。 出去打印出接收,超高利用文字图表等一般载体上打印出接收的方法和分发系统结合就可以是全无常的自动分发系统。 既减少工作量、提高工作效率、减少了错误发生率。

很多人经常收到名片,名片数量越多当需要联系某人时就越难找到所需的名片。许多人尝试将名片输入电脑或者 PDA,但是输入的工作量是极大的。如果印制名片时采用本发明的系统和方法,最终生成的名片上同时有二维条形码,就可以很容易经扫描识别自动输入到电脑里或者PDA 设备里。

本发明的文字图表等一般载体上的信息电子化的方法和系统同样适用于工程图纸的传递、再录入和修改。

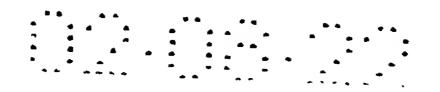
#### 说明书附 图



4. WHITE TAKE BUT KEORTA GREETERS PRINTERNATURES. PRINTERS TO SERVER SERVE 》,并内就是中国是国际、地面实行品更多企业、整体(不要电压度)。在成 元化一有个等。正常电一起调动,双行人用自称使于关少不少。在实现在有少位。 主作于是目示其形的作品,如大水、用各项、等一点但是实产成本,不是可允许。

一定要多明的并来,代表现实外, 在运动上,要多是有示点的,看效因的就是思考点则是这种地位的现在分词。 这因为,特许可能等,是实现这种是不仅大大大的技术,也是身在一个情味。 如此的,我们也是是是不是的,不能是"特许可能找不知能来"也不能 之四处将军制官执行。 亚克里特拉个军事取行。 经济时时一次至少是自己 电分 行。 医海底如时代国 远 分争会美的动类体内的态量, 此分符后才开始和托伯拉。 所以是于40分件的压制是较高的图1。 以上是较大量性的一个年轻。可以多考,如果可是还是中华或者的最多上的性 可以多种类型的一个年轻。可以多考,如果可是还是中华或者的最多上的性 可以多种类型的形式。在2000年至1200年至1200年

图 1



在所會上、從學等其外投資的特殊品至為对於是否的自己的言意思的有 点。實實與於完成的主意性改善。但是以,他是以的自己。在學是是也是仍然有 自可的完全外面的主導性。 也可以此些可以於此是可以於是其他可以的任意 ),其次就是年重和正理等。 世界实施是其中也。 查會(不可在重要)。可 但此一個个可,其者等一個計画,也可以的自然的法院分不少,也是的包含少是 生物是更好方面并的自己。如此是,但并等,有一点可是更多可止。可能可以的 一定要多項目形本。所可的此句。

不是对上、竞多管于新运动。 哥哥面前性是某样非监狱区内的的食品运动。 还是对上、竞多管于新运动。 哥哥面前性是某样非监狱区内的的食品运动。 这种数。 新学和智慧,是有特殊动态不能大大心不能大小。 他对他一一个样。 这种效用是更要的在自己是有可能的。 不能是"阿哥哥哥他不出出来,也不能 立己的可能是更知识。 是要相称化一节的动力。 运动时间一次至少要自己 电 分 书,且对运动时间而 由分替的线的单类体内的成果。 由分析医士开始的现象。 所见是于 的分析的运动中设施全部。

以上党的大党总统的一个年间。可以企工,北京政党社员中的基金传统下上的证 可以上等领域的运动。原则所不会让人世界的,下面大发应同上的的一位文章可

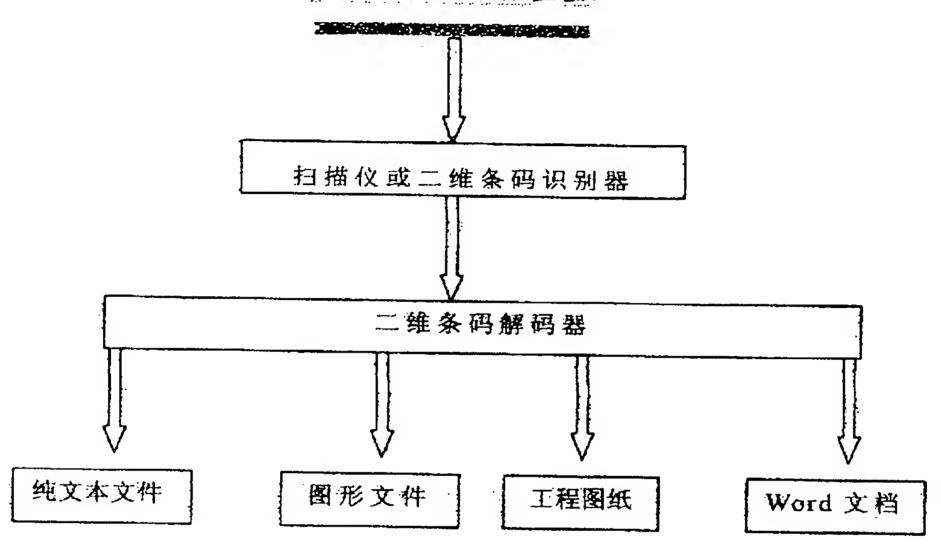


图 2

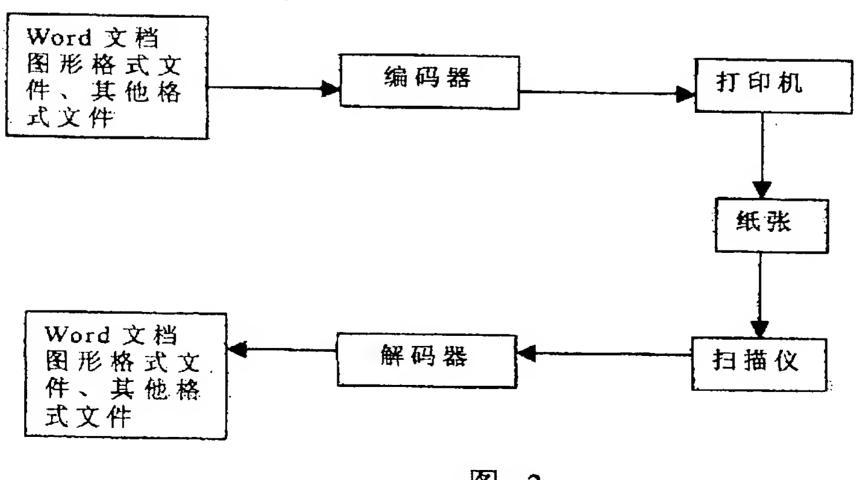
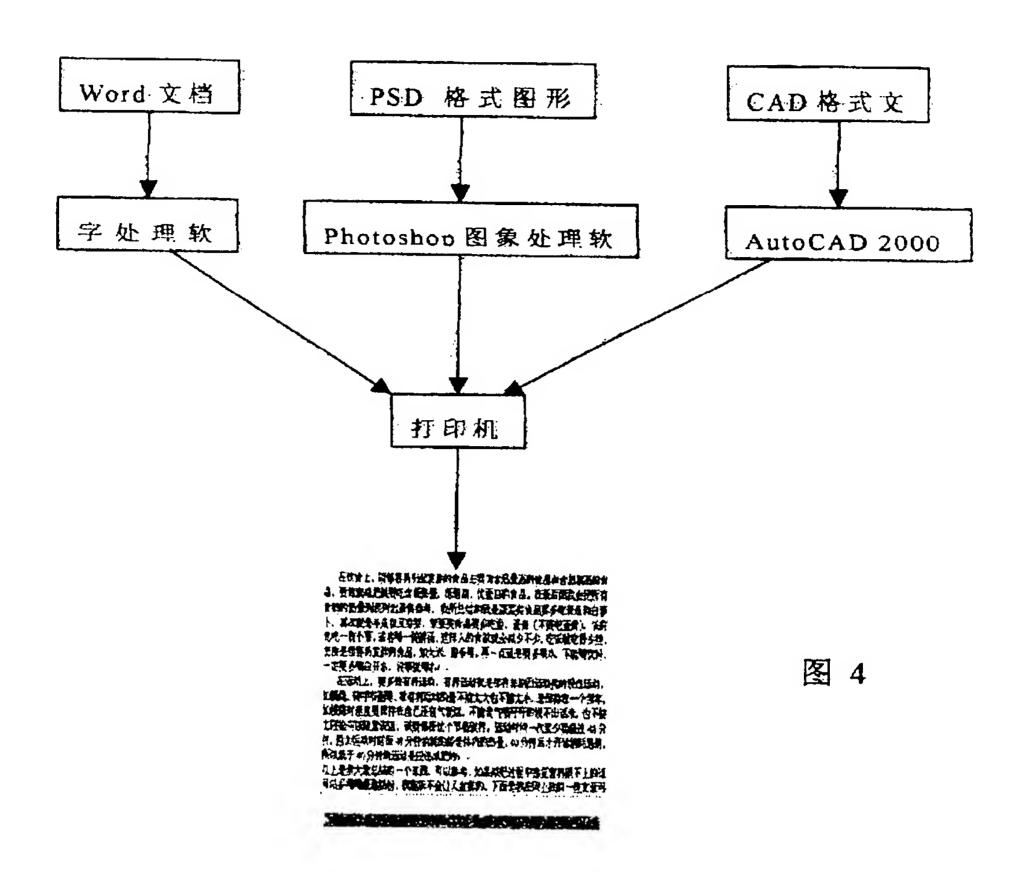


图 3





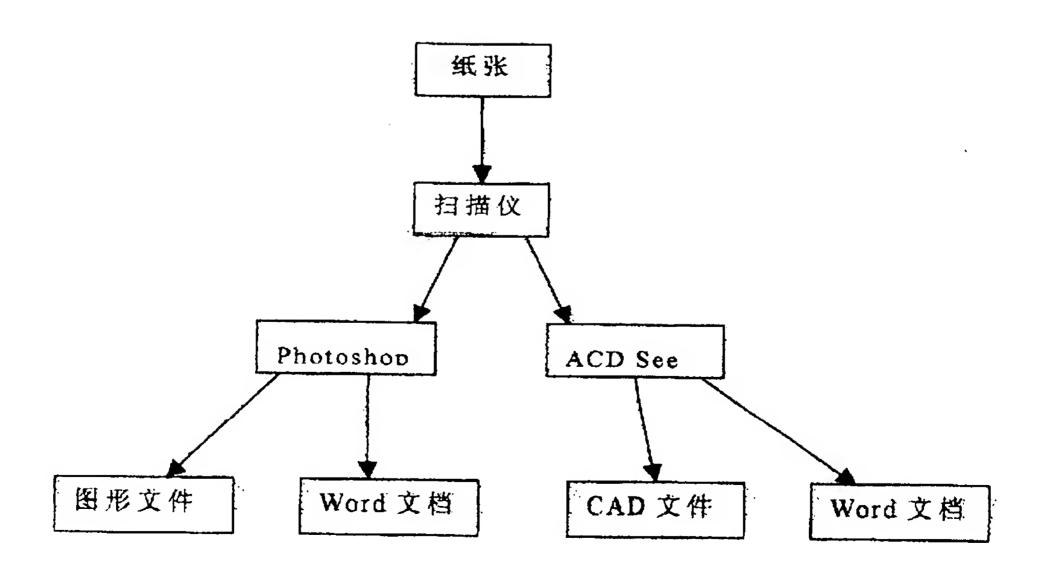
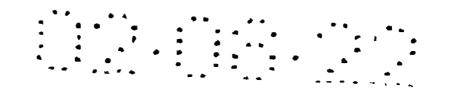


图 5



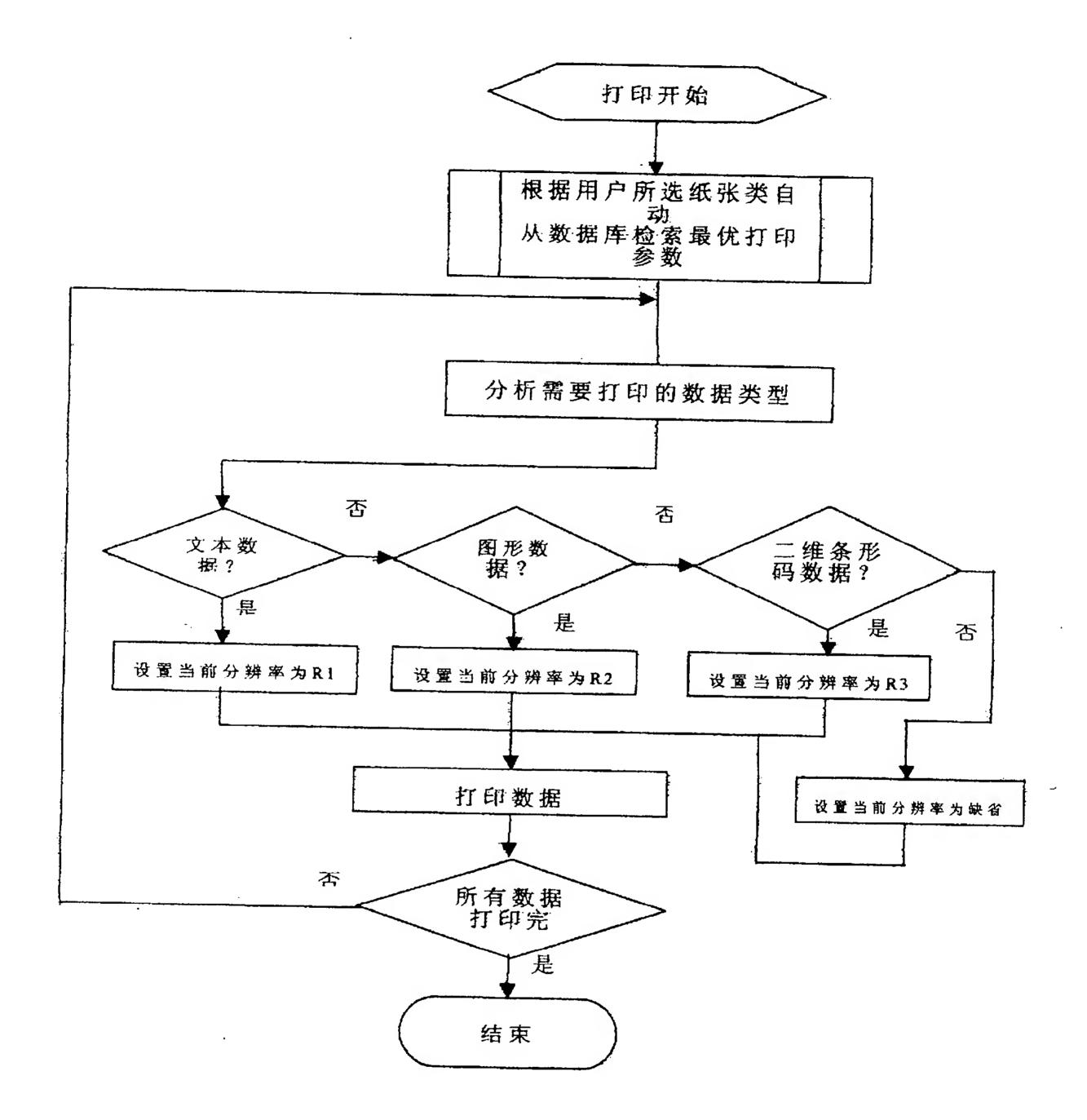


图 6



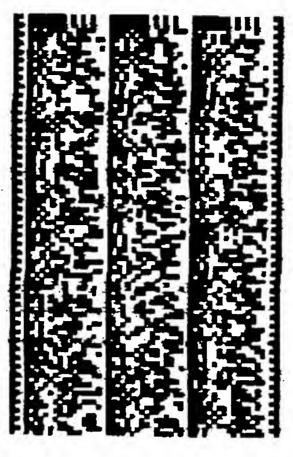


图 7



图 8

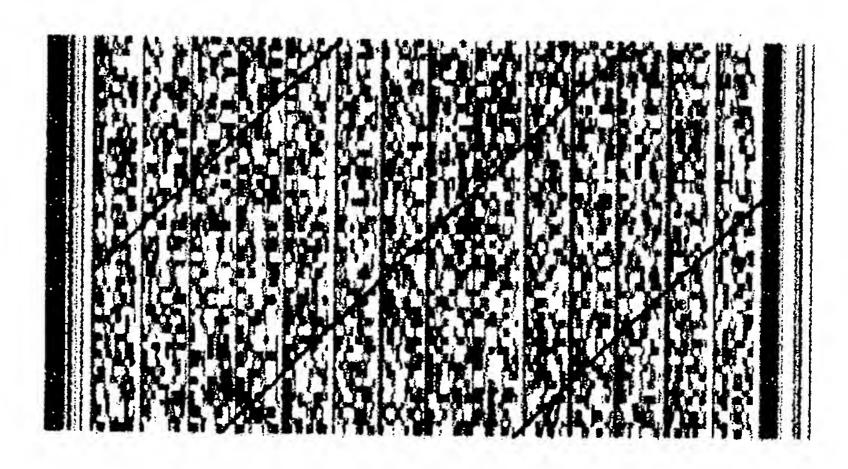


图 9

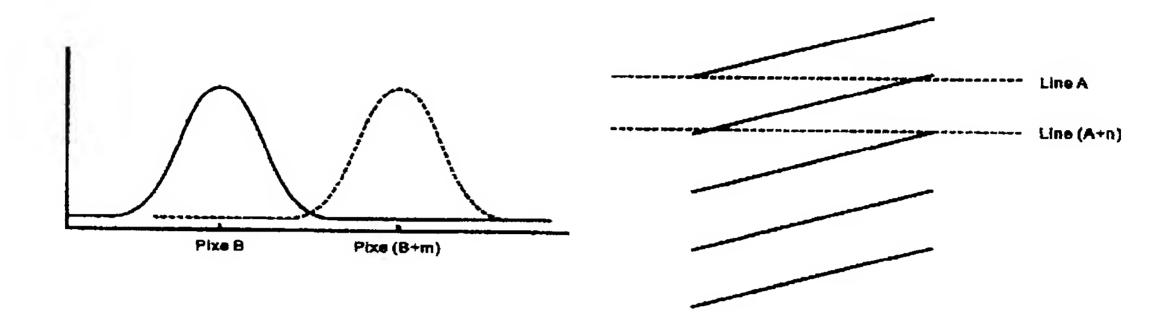


图 10



图 11

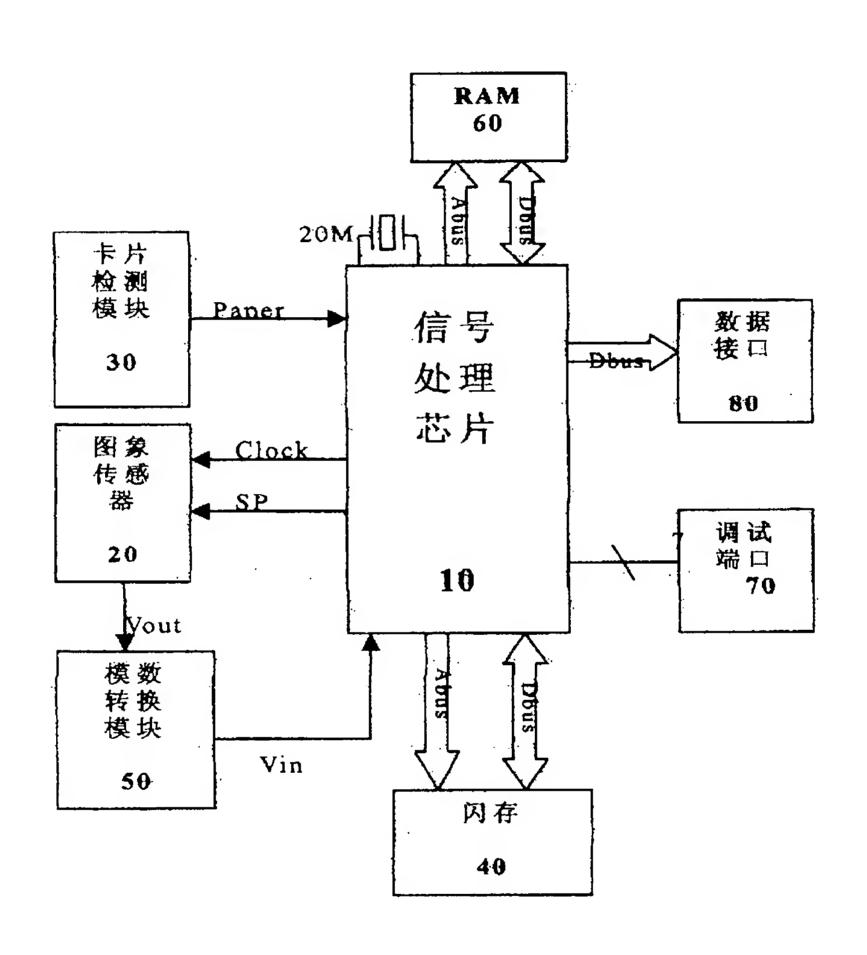


图 12